

AMENDMENTS TO THE CLAIMS

1. (currently amended) A method for maintaining records in a database comprising:

(a) receiving at least a collection of first data items and a collection of second data items~~[[,]] at least one of the first data items being associated with one of the second data items;~~

(b) disposing the first data items in a plurality of fields arranged in a predetermined format to form a first assemblage;

(c) disposing the second data items in a plurality of fields arranged in the predetermined format to form a second assemblage;

(d) modifying selected ones of the first data items and the second data items to conform to predetermined nomenclature;

(e) maintaining a record, in the database, having a plurality of fields arranged in the predetermined format; and

(f) determining whether a first particular data item in the predetermined nomenclature in a selected field of the first assemblage is identical to a second particular data item in the predetermined nomenclature in a field of the second assemblage corresponding to the selected field; and

~~(f)~~(g) if it is determined that the first particular data item in the predetermined nomenclature is identical to the second particular data item in the predetermined nomenclature, including in a field in the record a selected one of (1) the first data items in

the predetermined nomenclature and (2) the second data items in the predetermined nomenclature.

2. (original) The method of Claim 1 wherein at (b) and (c) the first data items and the second data items are separated, rearranged, or combined to form the first assemblage and the second assemblage respectively.

3. (currently amended) The method of Claim 1 wherein at ~~(f)~~(g) the first data items and the second data items are assigned accuracy rankings and are selected based upon their accuracy rankings.

4. (currently amended) A system for maintaining records in a database comprising:

a communications interface for receiving at least a collection of first data items and a collection of second data items[[,]] ~~at least one of the first data items being associated with one of the second data items;~~

a converter for disposing the first data items in a plurality of fields arranged in a predetermined format to form a first assemblage, and for disposing the second data items in a plurality of fields arranged in the predetermined format to form a second assemblage;

a normalizer for modifying selected ones of the first data items and the second data items to conform to predetermined nomenclature;

a database for maintaining a record, in the database, having a plurality of fields arranged in the predetermined format; and

a data masher for determining whether a first particular data item in the predetermined nomenclature in a selected field of the first assemblage is identical to a second particular data item in the predetermined nomenclature in a field of the second assemblage corresponding to the selected field, and for including in a field in the record a selected one of (1) the first data items in the predetermined nomenclature and (2) the second data items in the predetermined nomenclature, if it is determined that the first particular data item in the predetermined nomenclature is identical to the second particular data item in the predetermined nomenclature.

5. (currently amended) The system of Claim 4 wherein each of the ~~of the~~ first assemblage of data items and the second assemblage of data items in the predetermined nomenclature are associated with data accuracy rankings.

6. (currently amended) The system of Claim 5 wherein the data masher ~~selected~~ selects the first data items and the second data items based upon the data accuracy rankings.

7. (currently amended) A system for maintaining records in a database comprising:

a converter for converting at least a first collection of first data items and a second collection of second data items to at least a plurality of fields arranged in a predetermined format to form a first assemblage of data items and a second assemblage, respectively, the first and second assemblages each having data items in fields which are arranged in a predetermined format of data items, at least one of the first data items being associated with one of the second data items;

a normalizer for converting selected data items ~~ones~~ of the first assemblage of ~~data items~~ and the second assemblage of ~~data items~~ to conform to a predetermined nomenclature;

a database for maintaining a record~~[[,]] in the database[[,]]~~ having a plurality of fields arranged in the predetermined format; and

a data masher for determining a value representing a number of corresponding fields in the first assemblage and the second assemblage having identical data items therein, the data masher based on the value selecting at least one of the data items in the first assemblage and the second assemblage to form the record selecting at least one data item from the first assemblage of data items in the predetermined nomenclature, selecting data items from the second assemblage of data items in the predetermined nomenclature, and combining the at least one data items and the data items from the second assemblage to create the record.

Serial No. 09/992,987

8. (original) The system of Claim 7 wherein the converter separates, rearranges, or combines data to form the first assemblage of data items and the second assemblage of data items.

9. (original) The system of Claim 7 wherein each of the of the first assemblage of data items and the second assemblage of data items in the predetermined nomenclature are associated with data accuracy rankings.

10. (currently amended) The system of Claim 9 wherein the data masher ~~selected~~selects data items in the first assemblage and the second assemblage~~the first data items and the second data items~~ based upon the data accuracy ratings.

11. (currently amended) A system for maintaining records in a database comprising:

a converter for converting first records containing data items to uniform records having the data items organized in a uniform format, ~~a collection of first records having fields and data items into uniform records having fields and data items organized into a uniform format and where each of~~ wherein the uniform records have one or more of the data items conforming to a predetermined nomenclature and a balance of the one or more of the data items not conforming to the predetermined nomenclature;

a normalizer for converting the balance of the one or more data items to the predetermined nomenclature, and producing a collection of second records each having

selected data items organized in fields according to the uniform format, the selected data items conforming to the one or more of the data items that conform to a predetermined nomenclature and the balance of the one or more data items that are in the predetermined nomenclature; and

a data masher for selecting, from among the second records, at least a first selected record and a second selected record based on a value representing a number of corresponding fields in the first selected record and the second selected record having identical data items therein, and for selecting data items from the first selected record and second selected record to form a third data record, the third data record being equivalent records from the collection of second records and selecting data items from the equivalent records for combining into a third record having the selected data items arranged in the uniform format and stored in the database.

12. (original) The system of Claim 11 wherein the converter separates, rearranges, or combines the fields and the data items, from the first records, into the uniform format.

13. (original) The system of Claim 11 wherein before the normalizer converts the balance of the one or more data items it separates the one or more data items into components.

14. (currently amended) The system of Claim 11 wherein the data masher ~~selects equivalent records by comparing the data items in the second records on a field-by-field basis and determines a confidence level value of the second records based upon the similarities of the data items in the second records, and determining~~ determines that the ~~second records~~ first selected record and the second selected record are equivalent if the ~~confidence level values are~~ value is above a predetermined threshold level value.

15. (currently amended) Software recordable in a tangible medium which includes machine readable instructions for performing a process for building a database, which stores records corresponding to a plurality of data items, the process comprising:

(a) receiving at least a collection of first data items and a collection of second data items~~[[,]] at least one of the first data items being associated with one of the second data items;~~

(b) disposing the first data items in a plurality of fields arranged in a predetermined format to form a first assemblage;

(c) disposing the second data items in a plurality of fields arranged in the predetermined format to form a second assemblage;

(d) modifying selected ones of the first data items and the second data items to conform to predetermined nomenclature;

(e) maintaining a record, in the database, having a plurality of fields arranged in the predetermined format; and

(f) determining whether a first particular data item in the predetermined nomenclature in a selected field of the first assemblage is identical to a second particular data item in the predetermined nomenclature in a field of the second assemblage corresponding to the selected field; and

(f)(g) if it is determined that the first particular data item in the predetermined nomenclature is identical to the second particular data item in the predetermined nomenclature, including in a field in the record a selected one of (1) the first data items in the predetermined nomenclature and (2) the second data items in the predetermined nomenclature.

16. (original) The process of Claim 15 wherein at (b) and (c) the first data items and the second data items are separated, rearranged, or combined to form the first assemblage and the second assemblage respectively.

17. (currently amended) The process of Claim 15 wherein at ~~(f)~~(g) the first data items and the second data items are assigned accuracy rankings and are selected based upon their accuracy rankings.

18. (currently amended) A method for maintaining records in a database comprising:

converting first records containing data items to uniform records having the data items organized in a uniform format, a collection of first records having fields and data



~~items into uniform records having fields and data items and where each of wherein~~ the uniform records have one or more of the data items conforming to a predetermined nomenclature and a balance of the one or more data items not conforming to the predetermined nomenclature;

converting the balance of the one or more data items to the predetermined nomenclature[[],];

producing a collection of second records each having selected data items organized in fields according to the uniform format, the selected data items conforming to the ~~the one or more of the data items that conform to a predetermined nomenclature and the balance of the one or more data items that are in the~~ predetermined nomenclature; and

selecting, from among the second records, at least a first selected record and a second selected record based on a value representing a number of corresponding fields in the first selected record and the second selected record having identical data items therein; and

selecting data items from the first selected record and second selected record to form a third data record, the third data record being equivalent ~~records from the collection of second records and selecting the data items from the equivalent records for combining into a third record having the selected data items and stored in the database.~~

19. (original) The method of Claim 18 wherein the fields and the data items, from the first records, are separated, rearranged, or combined.

Serial No. 09/992,987

20. (original) The method of Claim 18 wherein the balance of the one or more data items are separated into components and converted into the predetermined nomenclature.

21. (currently amended) The method of Claim 18 wherein ~~equivalent records~~ are selected by comparing the data items in the second records on a field-by-field basis and determining a confidence level value of the second records based upon the similarities of the data items in the second records, and determining that the first selected record and the second selected record ~~second records~~ are determined to be equivalent if the ~~confidence level values are~~ value is above a predetermined threshold level value.